Most unambiguous loss-of-function *CPA1* mutations are unlikely to predispose to chronic pancreatitis

Jin-Huan Lin,[1,2,3] Arnaud Boulling,[1] Emmanuelle Masson,[1,4] David N. Cooper,[5] Zhao-Shen Li,[2,3] Claude Férec,[1,4] Zhuan Liao,[2,3] Jian-Min Chen[1]

[1]EFS, Univ Brest, Inserm, UMR 1078, GGB, F-29200 Brest, France

[2]Department of Gastroenterology, Changhai Hospital, The Second Military Medical University, Shanghai, China

[3]Shanghai Institute of Pancreatic Diseases, Shanghai, China

[4]CHU Brest, Service de Génétique, Brest, France

[5]Institute of Medical Genetics, School of Medicine, Cardiff University, Cardiff, United Kingdom

Word count (main text plus Figure and Table legends): 600

Abbreviations: CP, chronic pancreatitis; ER, endoplasmic reticulum; LoF, loss-of-function; NMD, nonsense-mediated RNA decay; PTC, premature termination codon

Keywords: chronic pancreatitis; *CPA1* gene; endoplasmic reticulum stress; loss-of-function variants; nonsense-mediated RNA decay

Correspondence to Dr Jian-Min Chen, INSERM UMR1078 – EFS – UBO, 22 avenue Camille Desmoulins, 29238 BREST, France; Tel: +33-2-98018174; Fax: +33-2-98016474; e-mail: jian-min.chen@univ-brest.fr

Prof. Zhuan Liao, Department of Gastroenterology, Changhai Hospital, The Second Military Medical University, 168 Changhai Road, Shanghai 200433, China; Tel: +86-21-31161004; Fax: 0086-21-55621735; e-mail: liaozhuan@smmu.edu.cn

We have read with interest the recent publication of Hegyi and Sahin-Tóth[1] reporting that chronic pancreatitis (CP)-predisposing *CPA1* mutations function through the misfolding pathway rather than through loss of CPA1 protein/activity. Herein, we explore an additional insight gleaned from this study beyond those discussed in an accompanying Editorial.[2].

In the original study reporting the association of *CPA1* variants with CP, all unambiguous loss-of-function (LoF) variants (e.g., nonsense mutations) were lumped together with missense mutations that functionally impaired the CPA1 protein.[3] However, unlike missense mutations, unambiguous LoF variants often result in transcripts that contain premature termination codons (PTC) and are thus prone to nonsense-mediated RNA decay (NMD). NMD detects and degrades PTC-containing transcripts, thereby preventing the accumulation of truncated proteins.[4,5] This implies that most unambiguous LoF *CPA1* variants would not be able to elicit ER stress and hence, in the light of the Hegyi and Sahin-Tóth study, will not predispose to CP. Indeed, the most frequently observed LoF variant, c.79C>T (p.Arg27*), was present at a lower frequency in European CP patients than in controls (Table 1). Additionally, we evaluated the pLI score for *CPA1* in the Genome Aggregation Database (genomAD; http://gnomad.broadinstitute.org/). The pLI score indicates the probability that a gene is intolerant to heterozygous LoF variants, ranging from 0 (completely tolerated) to 1.0 (extremely intolerant).[6] The pLI score for *CPA1* is 0. For the sake of comparison, *PRSS1* and *SPINK1* have pLI scores of 0 and 0.33 respectively; unambiguous LoF variants in the *PRSS1* gene actually protect against CP whereas unambiguous LoF variants in the *SPINK1* gene predispose to CP.[7]

Of the *CPA1* variants so far reported,[8] five may be regarded as unambiguous LoF variants by virtue of their mutation type and location (Table 1). We surveyed the clinical significance of these five LoF variants[8]. It would appear that only the classification of c.79C>T (p.Arg27*) as benign is strongly supported by genetic epidemiological data (Table 1). In

order to investigate the classifications of these five variants from a mechanistic standpoint, we tested whether the three coding variants that were predicted to generate PTCs, c.79C>T (p.Arg27*), c.357C>A (p.Tyr119*) and c.954_955delCA (p.Tys318*), would generate mutant transcripts that would be degraded by the NMD pathway. This was found to be the case for all three variants (figure 1). We also elucidated the precise splicing consequences of the two splice site mutations (i.e., c.148-1G>A and c.1072+1G>T) in a minigene system (see online supplementary figure S1); both gave rise to transcripts that were prone to NMD. All experimental details are described in online supplementary material, table S1 and figures S2 and S3.

In summary, prompted by the recent Hegyi and Sahin-Tóth study,[1] we provide evidence to suggest that most of the unambiguous LoF *CPA1* variants reported to date may not predispose to CP. Following the same line of reasoning, some of the previously characterized LoF *CPA1* missense mutations[2] may also not predispose to CP. In other words, the pathology may be confined to a small subset of *CPA1* mutations that are capable of eliciting ER stress. This may help to explain why rare functionally defective *CPA1* variants were not found to be associated with CP in a large Chinese cohort study.[9]

Contributors J-HL, AB and EM performed the functional assay. J-MC, ZL, Z-SL and CF designed the study. J-MC drafted the paper. DNC critically revised the manuscript. All authors analysed the data, contributed to revision of the manuscript and approved the final manuscript.

Funding J.H.L., a joint PhD student between the Changhai Hospital and INSERM U1078, was in receipt of a 20-month scholarship from the China Scholarship Council (No. 201706580018). Support for this study came from the Institut National de la Santé et de la

Recherche Médicale (INSERM), the Association des Pancréatites Chroniques Héréditaires, and the Association de Transfusion Sanguine et de Biogénétique Gaetan Saleun, France; the National Natural Science Foundation of China (81470884 and 81770636 (to Z.L.) and 81873588 (to Z.S.L)), the Chang Jiang Scholars Program of Ministry of Education (Q2015190 (to Z.L.)), and the Scientific Innovation Program of Shanghai Municipal Education Committee (to Z.L.), China.

Competing interests None declared.

REFERENCES

1. Hegyi E, Sahin-Tóth M. Human *CPA1* mutation causes digestive enzyme misfolding and chronic pancreatitis in mice. *Gut* 2019;68:301-12.
2. Rosendahl J. Unfolding the mystery of digestive enzyme misfolding in chronic pancreatitis. *Gut* 2019;68:181-2.
3. Witt H, Beer S, Rosendahl J, *et al*. Variants in *CPA1* are strongly associated with early onset chronic pancreatitis. *Nat Genet* 2013;45:1216-20.
4. Lykke-Andersen S, Jensen TH. Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. *Nat Rev Mol Cell Biol* 2015;16:665-77.
5. Goetz AE, Wilkinson M. Stress and the nonsense-mediated RNA decay pathway. *Cell Mol Life Sci* 2017;74:3509-31.
6. Lek M, Karczewski KJ, Minikel EV, *et al*. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 2016;536:285-91.
7. Chen JM, Férec C. Chronic pancreatitis: genetics and pathogenesis. *Annu Rev Genomics Hum Genet* 2009;10:63-87.
8. Sahin-Tóth M, Nemeth B. Genetic risk factors in chronic pancreatitis. http://www.pancreasgenetics.org/index.php (accessed 12 January 2019).
9. Wu H, Zhou DZ, Berki D, *et al*. No significant enrichment of rare functionally defective *CPA1* variants in a large Chinese idiopathic chronic pancreatitis cohort. *Hum Mutat* 2017;38:959-63.
10. Roberts NJ, Norris AL, Petersen GM, *et al*. Whole genome sequencing defines the genetic heterogeneity of familial pancreatic cancer. *Cancer Discov* 2016;6:166-75.

FIGURE LEGEND

Figure 1 Quantitative reverse transcriptase-PCR analysis of HEK293T cells transfected with *CPA1* cDNA expression constructs carrying respectively the wild-type and variant sequences. (A) mRNA expression levels of the three variant sequences relative to that of the wild-type sequence. (B) Relative mRNA expression levels of the variant sequences in transfected cells with (grey) and without (black) cycloheximide (an NMD inhibitor) treatment.

Table 1 The five unambiguous LoF *CPA1* variants discussed in this study

Region	Nucleotide change	Amino acid change	Cases (%) (*n*=1544)[a]	Controls (%) (*n*=6370)[b]	Current classification of clinical significance[8]	Functional analysis findings (this study)
Exon 2	c.79C>T	p.Arg27*	1 (0.06)	7 (0.11)	Benign	mRNA expression analysis demonstrated that the mutant transcript was subject to NMD.
Intron 2	c.148-1G>A	p.Leu50Hisfs*16 (previously termed p.Leu50_Glu127del)[2]	0 (0)	1 (0.02)	Likely pathogenic	Minigene splicing analysis revealed that the mutation primarily activated a cryptic 3'-splice site located 63 bp downstream of the normal one, resulting in the loss of the first 65 bp of exon 3 from the transcript. This would lead to a frameshift starting at amino acid position 50, with the new reading frame ending in a stop at position 16.
Exon 3	c.357C>A[c]	p.Tyr119*	–	–	Uncertain	mRNA expression analysis demonstrated that the mutant transcript was subject to NMD.
Exon 8	c.954_955delCA	p.Tyr318*	2 (0.13)	0 (0)	Pathogenic	mRNA expression analysis demonstrated that the mutant transcript was subject to NMD.
Intron 9	c.1072+1G>T	p.Asp330Ilefs*51	0 (0)	1 (0.02)	Likely pathogenic	Minigene splicing analysis confirmed that exon 9 was skipped. This would lead to a frame-shifting change starting at amino acid positon 330, with the new reading frame ending in a stop at position 51.

[a,b]Combined European data from reference 2 unless otherwise specified.
[c]Detected in a pancreatic cancer patient.[10]